

## SEQUENCE LISTING

### (1) GENERAL INFORMATION:

- (i) APPLICANT: Conneely, Orla M.  
Headon, Denis R.  
O'Malley, Bert W.  
May, Gregory S.
- (ii) TITLE OF INVENTION: Production of Recombinant Lactoferrin  
and Lactoferrin Polypeptides Using cDNA Sequences in  
Various Organisms
- (iii) NUMBER OF SEQUENCES: 14
- (iv) CORRESPONDENCE ADDRESS:
  - (A) ADDRESSEE: Pennie & Edmonds
  - (B) STREET: 1155 Avenue of the Americas
  - (C) CITY: New York
  - (D) STATE: New York
  - (E) COUNTRY: USA
  - (F) ZIP: 10036
- (v) COMPUTER READABLE FORM:
  - (A) MEDIUM TYPE: Floppy disk
  - (B) COMPUTER: IBM PC compatible
  - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
  - (D) SOFTWARE: PatentIn Release #1.0, Version #1.25
- (vi) CURRENT APPLICATION DATA:
  - (A) APPLICATION NUMBER: To be assigned
  - (B) FILING DATE: Concurrently herewith
  - (C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:
  - (A) APPLICATION NUMBER: 08/145,681
  - (B) FILING DATE: October 28, 1993
- (viii) ATTORNEY/AGENT INFORMATION:
  - (A) NAME: Albert P. Halluin
  - (B) REGISTRATION NUMBER: 25,227
  - (C) REFERENCE/DOCKET NUMBER: 8206-023
- (ix) TELECOMMUNICATION INFORMATION:
  - (A) TELEPHONE: 415-854-3660
  - (B) TELEFAX: 415-854-3694

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2360 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *H. sapiens*

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

GAATTCGAC CGCAGACATG AAACCTGTCT TCCTCGTCCT GCTGTTCTC GGGGCCCTCG	60
GACTGTGTCT GGCTGGCCGT AGGAGAAGGA GTGTTCACTG GTGCACCGTA TCCCAACCCG	120
AGGCCACAAA ATGCTTCCAA TGGCAAAGGA ATATGAGAAG AGTGCCTGGC CCTCCTGTCA	180
GCTGCATAAA GAGAGACTCC CCCATCCAGT GTATCCAGGC CATTGCGGAA AACAGGGCCG	240
ATGCTGTGAC CCTTGATGGT GGTTCATAT ACGAGGCAGG CCTGGCCCCC TACAACTGC	300
GACCTGTAGC GGCAGGAAGTC TACGGGACCG AAAGACAGCC ACGAACTCAC TATTATGCCG	360
TGGCTGTGGT GAAGAAGGGC GGCAGCTTTC AGCTGAACGA ACTGCAAGGT CTGAAGTCCT	420
GCCACACAGG CCTTCGCAGG ACCGCTGGAT GGAATGTGCC TATAGGGACA CTTCGTCCAT	480
TCTTGAATTG GACGGGTCCA CCTGAGCCCA TTGAGGCAGC TGTGGCCAGG TTCTTCTCAG	540
CCAGCTGTGT TCCCGGTGCA GATAAAGGAC AGTTCGCCAA CCTGTGTCGC CTGTGTGCGG	600
GGACAGGGGA AAACAAATGT GCCTTCTCCT CCCAGGAACC GTACTTCAGC TACTCTGGTG	660
CCTTCAAGTG TCTGAGAGAC GGGGCTGGAG ACGTGGCTTT TATCAGAGAG AGCACAGTGT	720
TTGAGGACCT GTCAGACGAG GCTGAAAGGG ACGAGTATGA GTTACTCTGC CCAGACAACA	780
CTCGGAAGCC AGTGGACAAG TTCAAAGACT GCCATCTGGC CCGGGTCCCT TCTCATGCCG	840
TTGTGGCACG AAGTGTGAAT GGCAAGGAGG ATGCCATCTG GAATCTTCTC CGCCAGGCAC	900
AGGAAAAGTT TGGAAAGGAC AAGTCACCGA AATTCCAGCT CTTTGGCTCC CCTAGTGGGC	960
AGAAAGATCT GCTGTTCAAG GACTCTGCCA TTGGGTTTTT GAGGGTGCCC CCGAGGATAG	1020
ATTCTGGGCT GTACCTTGGC TCCGGCTACT TCACTGCCAT CCAGAACTTG AGGAAAAGTG	1080
AGGAGGAAGT GGCTGCCCGG CGTGCGCGGG TCGTGTGGTG TGCGGTGGGC GAGCAGGAGC	1140
TGCGCAAGTG TAACCAAGTG AGTGGCTTGA GCGAAGGCAG CGTGACCTGC TCCTCGGCCT	1200
CCACCACAGA GGAAGTGCATC GCCCTGGTGC TGAAAGGAGA AGCTGATGCC ATGAGTTTGG	1260
ATGGAGGATA TGTGTACACT GCAGGCAAAT GTGGTTTGGT GCCTGTCCTG GCAGAGAACT	1320

ACAAATCCCA	ACAAAGCAGT	GACCCTGATC	CTAACTGTGT	GGATAGACCT	GTGGAAGGAT	1380
ATCTTGCTGT	GGCGGTGGTT	AGGAGATCAG	ACACTAGCCT	TACCTGGAAC	TCTGTGAAAG	1440
GCAAGAAGTC	CTGCCACACC	GCCGTGGACA	GGA CTGCAGG	CTGGAATATC	CCCATGGGCC	1500
TGCTCTTCAA	CCAGACGGGC	TCCTGCAAAT	TTGATGAATA	TTTCAGTCAA	AGCTGTGCCC	1560
CTGGGTCTGA	CCCAGATCT	AATCTCTGTG	CTCTGTGTAT	TGGCGACGAG	CAGGGTGAGA	1620
ATAAGTGCGT	GCCCAACAGC	AATGAGAGAT	ACTACGGCTA	CACTGGGGCT	TTCCGGTGCC	1680
TGGCTGAGAA	TGCTGGAGAC	GTTGCATTTG	TGAAAGATGT	CACTGTCTTG	CAGAACACTG	1740
ATGGAAATAA	CAATGAGGCA	TGGGCTAAGG	ATTTGAAGCT	GGCAGACTTT	GCGCTGCTGT	1800
GCCTCGATGG	CAAACGGAAG	CCTGTGACTG	AGGCTAGAAG	CTGCCATCTT	GCCATGGCCC	1860
CGAATCATGC	CGTGGTGTCT	CGGATGGATA	AGGTGGAACG	CCTGAAACAG	GTGCTGCTCC	1920
ACCAACAGGC	TAAATTTGGG	AGAAATGGAT	CTGACTGCCC	GGACAAGTTT	TGCTTATTCC	1980
AGTCTGAAAC	CAAAAACCTT	CTGTTCAATG	ACAACACTGA	GTGTCTGGCC	AGACTCCATG	2040
GCAAAACAAC	ATATGAAAAA	TATTTGGGAC	CACAGTATGT	CGCAGGCATT	ACTAATCTGA	2100
AAAAGTGCTC	AACCTCCCCC	CTCCTGGAAG	CCTGTGAATT	CCTCAGGAAG	TAAAACCGAA	2160
GAAGATGGCC	CAGCTCCCCA	AGAAAGCCTC	AGCCATTAC	TGCCCCCAGC	TCTTCTCCCC	2220
AGGTGTGTTG	GGGCCTTGGC	TCCCCTGCTG	AAGGTGGGGA	TTGCCCATCC	ATCTGCTTAC	2280
AATTCCTGCTG	TGTCGTCTTA	GCAAGAAGTA	AAATGAGAAA	TTTTGTTGAA	AAAAAAAAAA	2340
AAAAAAAAAA	AAAAAAAAAA					2360

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 711 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: H. sapiens

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met	Lys	Leu	Val	Phe	Leu	Val	Leu	Leu	Phe	Leu	Gly	Ala	Leu	Gly	Leu	
1				5					10					15		
Cys	Leu	Ala	Gly	Arg	Arg	Arg	Arg	Ser	Val	Gln	Trp	Cys	Thr	Val	Ser	
			20					25					30			
Gln	Pro	Glu	Ala	Thr	Lys	Cys	Phe	Gln	Trp	Gln	Arg	Asn	Met	Arg	Arg	
		35					40					45				
Val	Arg	Gly	Pro	Pro	Val	Ser	Cys	Ile	Lys	Arg	Asp	Ser	Pro	Ile	Gln	
	50					55					60					
Cys	Ile	Gln	Ala	Ile	Ala	Glu	Asn	Arg	Ala	Asp	Ala	Val	Thr	Leu	Asp	
65					70					75					80	
Gly	Gly	Phe	Ile	Tyr	Glu	Ala	Gly	Leu	Ala	Pro	Tyr	Lys	Leu	Arg	Pro	
			85					90						95		
Val	Ala	Ala	Glu	Val	Tyr	Gly	Thr	Glu	Arg	Gln	Pro	Arg	Thr	His	Tyr	
			100					105						110		
Tyr	Ala	Val	Ala	Val	Val	Lys	Lys	Gly	Gly	Ser	Phe	Gln	Leu	Asn	Glu	
		115					120					125				
Leu	Gln	Gly	Leu	Lys	Ser	Cys	His	Thr	Gly	Leu	Arg	Arg	Thr	Ala	Gly	
		130				135						140				
Trp	Asn	Val	Pro	Ile	Gly	Thr	Leu	Arg	Pro	Phe	Leu	Asn	Trp	Thr	Gly	
145					150					155					160	
Pro	Pro	Glu	Pro	Ile	Glu	Ala	Ala	Val	Ala	Arg	Phe	Phe	Ser	Ala	Ser	
				165					170					175		
Cys	Val	Pro	Gly	Ala	Asp	Lys	Gly	Gln	Phe	Pro	Asn	Leu	Cys	Arg	Leu	
			180					185					190			
Cys	Ala	Gly	Thr	Gly	Glu	Asn	Lys	Cys	Ala	Phe	Ser	Ser	Gln	Glu	Pro	
		195					200					205				
Tyr	Phe	Ser	Tyr	Ser	Gly	Ala	Phe	Lys	Cys	Leu	Arg	Asp	Gly	Ala	Gly	
	210					215						220				
Asp	Val	Ala	Phe	Ile	Arg	Glu	Ser	Thr	Val	Phe	Glu	Asp	Leu	Ser	Asp	
225					230					235					240	

Glu Ala Glu Arg Asp Glu Tyr Glu Leu Leu Cys Pro Asp Asn Thr Arg  
 245 250 255  
 Lys Pro Val Asp Lys Phe Lys Asp Cys His Leu Ala Arg Val Pro Ser  
 260 265 270  
 His Ala Val Val Ala Arg Ser Val Asn Gly Lys Glu Asp Ala Ile Trp  
 275 280 285  
 Asn Leu Leu Arg Gln Ala Gln Glu Lys Phe Gly Lys Asp Lys Ser Pro  
 290 295 300  
 Lys Phe Gln Leu Phe Gly Ser Pro Ser Gly Gln Lys Asp Leu Leu Phe  
 305 310 315 320  
 Lys Asp Ser Ala Ile Gly Phe Ser Arg Val Pro Pro Arg Ile Asp Ser  
 325 330 335  
 Gly Leu Tyr Leu Gly Ser Gly Tyr Phe Thr Ala Ile Gln Asn Leu Arg  
 340 345 350  
 Lys Ser Glu Glu Glu Val Ala Ala Arg Arg Ala Arg Val Val Trp Cys  
 355 360 365  
 Ala Val Gly Glu Gln Glu Leu Arg Lys Cys Asn Gln Trp Ser Gly Leu  
 370 375 380  
 Ser Glu Gly Ser Val Thr Cys Ser Ser Ala Ser Thr Thr Glu Asp Cys  
 385 390 395 400  
 Ile Ala Leu Val Leu Lys Gly Glu Ala Asp Ala Met Ser Leu Asp Gly  
 405 410 415  
 Gly Tyr Val Tyr Thr Ala Gly Lys Cys Gly Leu Val Pro Val Leu Ala  
 420 425 430  
 Glu Asn Tyr Lys Ser Gln Gln Ser Ser Asp Pro Asp Pro Asn Cys Val  
 435 440 445  
 Asp Arg Pro Val Glu Gly Tyr Leu Ala Val Ala Val Val Arg Arg Ser  
 450 455 460  
 Asp Thr Ser Leu Thr Trp Asn Ser Val Lys Gly Lys Lys Ser Cys His  
 465 470 475 480  
 Thr Ala Val Asp Arg Thr Ala Gly Trp Asn Ile Pro Met Gly Leu Leu  
 485 490 495  
 Phe Asn Gln Thr Gly Ser Cys Lys Phe Asp Glu Tyr Phe Ser Gln Ser  
 500 505 510  
 Cys Ala Pro Gly Ser Asp Pro Arg Ser Asn Leu Cys Ala Leu Cys Ile  
 515 520 525  
 Gly Asp Glu Gln Gly Glu Asn Lys Cys Val Pro Asn Ser Asn Glu Arg  
 530 535 540  
 Tyr Tyr Gly Tyr Thr Gly Ala Phe Arg Cys Leu Ala Glu Asn Ala Gly  
 545 550 555 560  
 Asp Val Ala Phe Val Lys Asp Val Thr Val Leu Gln Asn Thr Asp Gly  
 565 570 575

Asn	Asn	Asn	Glu	Ala	Trp	Ala	Lys	Asp	Leu	Lys	Leu	Ala	Asp	Phe	Ala		
			580					585					590				
Leu	Leu	Cys	Leu	Asp	Gly	Lys	Arg	Lys	Pro	Val	Thr	Glu	Ala	Arg	Ser		
		595					600					605					
Cys	His	Leu	Ala	Met	Ala	Pro	Asn	His	Ala	Val	Val	Ser	Arg	Met	Asp		
	610					615					620						
Lys	Val	Glu	Arg	Leu	Lys	Gln	Val	Leu	Leu	His	Gln	Gln	Ala	Lys	Phe		
625					630					635					640		
Gly	Arg	Asn	Gly	Ser	Asp	Cys	Pro	Asp	Lys	Phe	Cys	Leu	Phe	Gln	Ser		
			645						650					655			
Glu	Thr	Lys	Asn	Leu	Leu	Phe	Asn	Asp	Asn	Thr	Glu	Cys	Leu	Ala	Arg		
			660					665					670				
Leu	His	Gly	Lys	Thr	Thr	Tyr	Glu	Lys	Tyr	Leu	Gly	Pro	Gln	Tyr	Val		
		675					680					685					
Ala	Gly	Ile	Thr	Asn	Leu	Lys	Lys	Cys	Ser	Thr	Ser	Pro	Leu	Leu	Glu		
	690					695					700						
Ala	Cys	Glu	Phe	Leu	Arg	Lys											
705					710												

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2347 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Bos taurus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

GAGCCTTCGT TCCGGAGTCG CCCAGGACG CCAGCCCATG AAGCTCTTCG TCCCCGCCCT	60
CCTGTCCCTT GGAGCCCTTG GACTGTGTCT GGCTGCCCCG AGGAAAACG TTCGATGGTG	120
TACCATCTCC CAACCTGAGT GGTTCAAATG CCGCAGATGG CAGTGGAGGA TGAAGAAGCT	180
GGGTGCTCCC TCTATCACCT GTGTGAGGCG GGCCTTTGCC TTGGAATGTA TTCCGGGCAT	240
CGCGGAGAAA AAGGCGGATG CTGTGACCCT GGATGGTGGC ATGGTGTTTG AGGCGGGCCG	300
GGACCCCTAC AAAGTGCAGC CAGTAGCAGC AGAGATCTAT GGGACGAAAG AGTCTCCCCA	360
AACCCACTAT TATGCTGTGG CCGTCGTGAA GAAGGGCAGC AACTTTCAGC TGGACCAGCT	420
GCAAGGCCGG AAGTCCTGCC ATACGGGCCT TGGCAGGTCC GCTGGGTGGA TCATCCCTAT	480
GGGAATCCTT CGCCCGTACT TGAGCTGGAC AGAGTCACTC GAGCCCCTCC AGGGAGCTGT	540
GGCTAAATTC TTCTCTGCCA GCTGTGTTCC CTGCATTGAT AGACAAGCAT ACCCCAACCT	600
GTGTCAACTG TGCAAGGGGG AGGGGGAGAA CCAGTGTGCC TGCTCCTCCC GGGAACCATA	660
CTTCGGTTAT TCTGGTGCCT TCAAGTGTCT GCAGGACGGG GCTGGAGACG TGGCTTTTGT	720
TAAAGAGACG ACAGTGTTTG AGAACTTGCC AGAGAAGGCT GACAGGGACC AGTATGAGCT	780
TCTCTGCCTG AACAACAGTC GGGCGCCAGT GGATGCGTTC AAGGAGTGCC ACCTGGCCCA	840
GGTCCCTTCT CATGCTGTCTG TGGCCCGAAG TGTGGATGGC AAGGAAGACT TGATCTGGAA	900
GCTTCTCAGC AAGGCGCAGG AGAAATCTGG AAAAAACAAG TCTCGGAGCT TCCAGCTCTT	960
TGGCTCTCCA CCCGGCCAGA GGGACCTGCT GTTCAAAGAC TCTGCTCTTG GGTTTTTGAG	1020
GATCCCTCG AAGGTAGATT CGGCGCTGTA CCTGGGCTCC CGCTACTTGA CCACCTTGAA	1080
GAACCTCAGG GAAACTGCGG AGGAGGTGAA GGC GCGGTAC ACCAGGGTCG TGTGGTGTGC	1140
CGTGGGACCT GAGGAGCAGA AGAAGTGCCA GCAGTGGAGC CAGCAGAGCG GCCAGAACGT	1200
GACCTGTGCC ACGGCGTCCA CCACTGACGA CTGCATCGTC CTGGTGCTGA AAGGGGAAGC	1260
AGATGCCCTG AACTTGGATG GAGGATATAT CTACACTGCG GGCAAGTGTG GCCTGGTGCC	1320

TGTCCTGGCA	GAGAACCGGA	AATCCTCCAA	ACACAGTAGC	CTAGATTGTG	TGCTGAGACC	1380
AACGGAAGGG	TACCTTGCCG	TGGCAGTTGT	CAAGAAAGCA	AATGAGGGGC	TCACATGGAA	1440
TTCTCTGAAA	GACAAGAAGT	CGTGCCACAC	CGCCGTGGAC	AGGACTGCAG	GCTGGAACAT	1500
CCCCATGGGC	CTGATCGTCA	ACCAGACAGG	CTCCTGCGCA	TTTGATGAAT	TCTTTAGTCA	1560
GAGCTGTGCC	CCTGGGGCTG	ACCCGAAATC	CAGACTCTGT	GCCTTGTGTG	CTGGCGATGA	1620
CCAGGGCCTG	GACAAGTGTG	TGCCCAACTC	TAAGGAGAAG	TACTATGGCT	ATACCGGGGC	1680
TTTCAGGTGC	CTGGCTGAGG	ACGTTGGGGA	CGTTGCCTTT	GTGAAAAACG	ACACAGTCTG	1740
GGAGAACACG	AATGGAGAGA	GCACTGCAGA	CTGGGCTAAG	AACTTGAATC	GTGAGGACTT	1800
CAGGTTGCTC	TGCCTCGATG	GCACCAGGAA	GCCTGTGACG	GAGGCTCAGA	GCTGCCACCT	1860
GGCGGTGGCC	CCGAATCACG	CTGTGGTGTC	TCGGAGCGAT	AGGGCAGCAC	ACGTGAAACA	1920
GGTGCTGCTC	CACCAGCAGG	CTCTGTTTGG	GAAAAATGGA	AAAAACTGCC	CGGACAAGTT	1980
TTGTTTGTTT	AAATCTGAAA	CCAAAAACCT	TCTGTTCAAT	GACAACACTG	AGTGTCTGGC	2040
CAAACCTGGA	GGCAGACCAA	CGTATGAAGA	ATATTGGGG	ACAGAGTATG	TCACGGCCAT	2100
TGCCAACCTG	AAAAAATGCT	CAACCTCCCC	GCTTCTGGAA	GCCTGCGCCT	TCCTGACGAG	2160
GTAAAGCCTG	CAAAGAAGCT	AGCCTGCCTC	CCTGGGCCTC	AGCTCCTCCC	TGCTCTCAGC	2220
CCCAATCTCC	AGGCGCGAGG	GACCTTCCTC	TCCCTTCCTG	AAGTCGGATT	TTTGCCAAGC	2280
TCATCAGTAT	TTACAATTCC	CTGCTGTCAT	TTTAGCAAGA	AATAAAATTA	GAAATGCTGT	2340
TGAAAAA						2347



(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 708 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Bos taurus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met	Lys	Leu	Phe	Val	Pro	Ala	Leu	Leu	Ser	Leu	Gly	Ala	Leu	Gly	Leu	
1				5					10					15		
Cys	Leu	Ala	Ala	Pro	Arg	Lys	Asn	Val	Arg	Trp	Cys	Thr	Ile	Ser	Gln	
			20				25						30			
Pro	Glu	Trp	Phe	Lys	Cys	Arg	Arg	Trp	Gln	Trp	Arg	Met	Lys	Lys	Leu	
		35				40						45				
Gly	Ala	Pro	Ser	Ile	Thr	Cys	Val	Arg	Arg	Ala	Phe	Ala	Leu	Glu	Cys	
	50					55					60					
Ile	Pro	Gly	Ile	Ala	Glu	Lys	Lys	Ala	Asp	Ala	Val	Thr	Leu	Asp	Gly	
65					70				75					80		
Gly	Met	Val	Phe	Glu	Ala	Gly	Arg	Asp	Pro	Tyr	Lys	Leu	Arg	Pro	Val	
			85						90					95		
Ala	Ala	Glu	Ile	Tyr	Gly	Thr	Lys	Glu	Ser	Pro	Gln	Thr	His	Tyr	Tyr	
			100					105					110			
Ala	Val	Ala	Val	Val	Lys	Lys	Gly	Ser	Asn	Phe	Gln	Leu	Asp	Gln	Leu	
		115					120					125				
Gln	Gly	Arg	Lys	Ser	Cys	His	Thr	Gly	Leu	Gly	Arg	Ser	Ala	Gly	Trp	
	130					135					140					
Ile	Ile	Pro	Met	Gly	Ile	Leu	Arg	Pro	Tyr	Leu	Ser	Trp	Thr	Glu	Ser	
145					150					155					160	
Leu	Glu	Pro	Leu	Gln	Gly	Ala	Val	Ala	Lys	Phe	Phe	Ser	Ala	Ser	Cys	
			165						170					175		
Val	Pro	Cys	Ile	Asp	Arg	Gln	Ala	Tyr	Pro	Asn	Leu	Cys	Gln	Leu	Cys	
		180					185						190			
Lys	Gly	Glu	Gly	Glu	Asn	Gln	Cys	Ala	Cys	Ser	Ser	Arg	Glu	Pro	Tyr	
	195					200						205				
Phe	Gly	Tyr	Ser	Gly	Ala	Phe	Lys	Cys	Leu	Gln	Asp	Gly	Ala	Gly	Asp	
	210				215						220					
Val	Ala	Phe	Val	Lys	Glu	Thr	Thr	Val	Phe	Glu	Asn	Leu	Pro	Glu	Lys	
225				230					235						240	

Ala Asp Arg Asp Gln Tyr Glu Leu Leu Cys Leu Asn Asn Ser Arg Ala  
 245 250 255  
 Pro Val Asp Ala Phe Lys Glu Cys His Leu Ala Gln Val Pro Ser His  
 260 265 270  
 Ala Val Val Ala Arg Ser Val Asp Gly Lys Glu Asp Leu Ile Trp Lys  
 275 280 285  
 Leu Leu Ser Lys Ala Gln Glu Lys Ser Gly Lys Asn Lys Ser Arg Ser  
 290 295 300  
 Phe Gln Leu Phe Gly Ser Pro Pro Gly Gln Arg Asp Leu Leu Phe Lys  
 305 310 315 320  
 Asp Ser Ala Leu Gly Phe Leu Arg Ile Pro Ser Lys Val Asp Ser Ala  
 325 330 335  
 Leu Tyr Leu Gly Ser Arg Tyr Leu Thr Thr Leu Lys Asn Leu Arg Glu  
 340 345 350  
 Thr Ala Glu Glu Val Lys Ala Arg Tyr Thr Arg Val Val Trp Cys Ala  
 355 360 365  
 Val Gly Pro Glu Glu Gln Lys Lys Cys Gln Gln Trp Ser Gln Gln Ser  
 370 375 380  
 Gly Gln Asn Val Thr Cys Ala Thr Ala Ser Thr Thr Asp Asp Cys Ile  
 385 390 395 400  
 Val Leu Val Leu Lys Gly Glu Ala Asp Ala Leu Asn Leu Asp Gly Gly  
 405 410 415  
 Tyr Ile Tyr Thr Ala Gly Lys Cys Gly Leu Val Pro Val Leu Ala Glu  
 420 425 430  
 Asn Arg Lys Ser Ser Lys His Ser Ser Leu Asp Cys Val Leu Arg Pro  
 435 440 445  
 Thr Glu Gly Tyr Leu Ala Val Ala Val Val Lys Lys Ala Asn Glu Gly  
 450 455 460  
 Leu Thr Trp Asn Ser Leu Lys Asp Lys Lys Ser Cys His Thr Ala Val  
 465 470 475 480  
 Asp Arg Thr Ala Gly Trp Asn Ile Pro Met Gly Leu Ile Val Asn Gln  
 485 490 495  
 Thr Gly Ser Cys Ala Phe Asp Glu Phe Phe Ser Gln Ser Cys Ala Pro  
 500 505 510  
 Gly Ala Asp Pro Lys Ser Arg Leu Cys Ala Leu Cys Ala Gly Asp Asp  
 515 520 525  
 Gln Gly Leu Asp Lys Cys Val Pro Asn Ser Lys Glu Lys Tyr Tyr Gly  
 530 535 540  
 Tyr Thr Gly Ala Phe Arg Cys Leu Ala Glu Asp Val Gly Asp Val Ala  
 545 550 555 560  
 Phe Val Lys Asn Asp Thr Val Trp Glu Asn Thr Asn Gly Glu Ser Thr  
 565 570 575

Ala	Asp	Trp	Ala	Lys	Asn	Leu	Asn	Arg	Glu	Asp	Phe	Arg	Leu	Leu	Cys	
			580					585					590			
Leu	Asp	Gly	Thr	Arg	Lys	Pro	Val	Thr	Glu	Ala	Gln	Ser	Cys	His	Leu	
		595					600					605				
Ala	Val	Ala	Pro	Asn	His	Ala	Val	Val	Ser	Arg	Ser	Asp	Arg	Ala	Ala	
		610				615					620					
His	Val	Lys	Gln	Val	Leu	Leu	His	Gln	Gln	Ala	Leu	Phe	Gly	Lys	Asn	
					630					635					640	
Gly	Lys	Asn	Cys	Pro	Asp	Lys	Phe	Cys	Leu	Phe	Lys	Ser	Glu	Thr	Lys	
				645					650						655	
Asn	Leu	Leu	Phe	Asn	Asp	Asn	Thr	Glu	Cys	Leu	Ala	Lys	Leu	Gly	Gly	
			660					665						670		
Arg	Pro	Thr	Tyr	Glu	Glu	Tyr	Leu	Gly	Thr	Glu	Tyr	Val	Thr	Ala	Ile	
			675				680						685			
Ala	Asn	Leu	Lys	Lys	Cys	Ser	Thr	Ser	Pro	Leu	Leu	Glu	Ala	Cys	Ala	
			690			695					700					
Phe	Leu	Thr	Arg													
			705													

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2259 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Sus scrofa*

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

ACATGAAGCT CTTCATCCCC GCCCTGCTGT TCCTCGGGAC ACTTGGACTG TGTCTGGCTG	60
CCCCTAAGAA AGGGGTTCGA TGGTGTGTCA TATCCACAGC AGAGTATTCA AAATGCCGCC	120
AGTGGCAATC AAAGATAAGA AGAACTAATC CCATGTTCTG CATAAGGAGG GCTTCTCCCA	180
CTGACTGTAT CCGGGCCATC GCGGCAAAA GGCAGATGC TGTGACCCTT GATGGTGGTT	240
TGGTGTTTGA AGCAGACCAG TACAACTGC GGCCGGTAGC AGCGGAGATC TACGGGACAG	300
AAGAGAATCC CCAAACCTAC TATTATGCTG TGGCTGTAGT GAAGAAAGGT TTCAACTTTC	360
AGAACCAGCT ACAAGGTCGA AAGTCCTGCC ACACAGGCCT TGGCAGGTCT GCCGGGTGGA	420
ATATCCCTAT AGGGTTACTT CGCCGGTTCT TGGACTGGGC AGGGCCACCT GAGCCCCCTC	480
AGAAAGCTGT GGCCAAATTC TTCTCTCAGA GCTGTGTGCC CTGCGCAGAT GGAAATGCGT	540
ATCCCAACCT GTGTCAGCTG TGCATAGGGA AAGGGAAAGA TAAATGTGCT TGTTCCTCCC	600
AGGAACCGTA TTTTGGCTAT TCCGGTGCCT TCAACTGTCT GCACAAAGGG ATTGGAGATG	660
TGGCTTTTGT CAAGGAGAGT ACAGTGTTTG AGAACCTGCC ACAGAAGGCT GACCGGGACA	720
AATACGAGCT ACTCTGCCCC GACAATACTC GAAAGCCAGT GGAAGCATTC AGGGAGTGCC	780
ACCTTGCCCCG GGTCCCTTCT CATGCTGTTG TGGCCCGAAG TGTGAATGGC AAGGAGAACT	840
CCATCTGGGA GCTTCTCTAC CAGTCACAGA AAAAGTTTGG AAAAAGCAAT CCACAGGAGT	900
TCCAGCTCTT TGGCTCTCCT GGTCAGCAGA AGGACCTCCT GTTTAGAGAT GCTACCATCG	960
GGTTTTTTGAA GATCCCCTCA AAGATAGATT CTAAGCTGTA CCTGGGCCTC CCGTACCTTA	1020
CTGCCATCCA GGGCCTGAGG GAAACGGCAG CGGAGGTGGA GGCGCGGCAG GCGAAGGTCG	1080
TGTGGTGCGC CGTGGGTCCA GAGGAGCTGC GCAAGTGCCG GCAGTGGAGC AGCCAGAGCA	1140
GCCAGAACCT GAACTGCAGC CTGGCCTCCA CCACCGAGGA CTGCATCGTC CAGGTGCTGA	1200
AAGGAGAAGC TGATGCTATG AGCTTGGATG GAGGATTTAT CTACACTGCG GGCAAGTGTG	1260
GTTTGGTGCC TGTCTTGCA GAGAACC AAAAATCTCGCCA AAGCAGTAGC TCAGACTGTG	1320

TGCATAGACC	AACACAAGGG	TATTTTGCCG	TGGCGGTTGT	CAGGAAAGCA	AATGGTGGTA	1380
TCACCTGGAA	CTCTGTGAGA	GGCACGAAGT	CCTGCCACAC	TGCTGTGGAC	AGGACAGCAG	1440
GCTGGAACAT	CCCCATGGGC	CTGCTTGTC	ACCAGACAGG	CTCCTGCAA	TTTGACGAAT	1500
TCTTTAGTCA	AAGCTGTGCT	CCTGGGTCTC	AGCCGGGATC	CAATCTCTGT	GCACTGTGTG	1560
TTGGCAATGA	CCAGGGCGTG	GACAAGTGTG	TGCCCAACAG	TAATGAGAGA	TACTATGGTT	1620
ACACCGGGGC	TTTCAGGTGC	CTGGCTGAGA	ATGCTGGGGA	TGTGGCGTTT	GTGAAAGATG	1680
TCACTGTCTT	GGACAACACG	AATGGACAGA	ACACAGAAGA	GTGGGCCAGG	GAATTGAGGT	1740
CAGATGACTT	TGAGCTGCTG	TGCCTTGATG	GCACCAGGAA	GCCTGTGACT	GAGGCTCAGA	1800
ACTGTCACCT	GGCTGTGGCC	CCCAGTCATG	CTGTGGTCTC	TCGGAAGGAA	AAGGCAGCAC	1860
AGGTGGAACA	GGTGCTACTC	ACTGAGCAGG	CTCAGTTTGG	AAGATACGGA	AAAGACTGCC	1920
CGGACAAGTT	TTGCTTGTTT	CGGTCTGAGA	CCAAAAACCT	TCTGTTCAAC	GACAACACGG	1980
AGGTTCTGGC	CCAACCTCAA	GGCAAAACAA	CATACGAAAA	ATATTTGGGA	TCAGAGTATG	2040
TCACAGCCAT	CGCTAACCTG	AAACAGTGCT	CAGTCTCCCC	GCTTCTGGAA	GCCTGTGCCT	2100
TCATGATGAG	GTAAAACCGG	AAAAGAAGCT	GCCCGCCTCC	CCAGGGGCCT	CAGCTTTCCC	2160
TCCTCCCGTC	TTGATTCCCA	GCTGCCCTGG	GCCTGCCTCT	CTCCCTTCCT	GAGGGCAGAC	2220
TTTGTTCAGC	TCATCCGTTT	TCACAATTCC	CTCGTGCCG			2259

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 703 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Sus scrofa*

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Met	Lys	Leu	Phe	Ile	Pro	Ala	Leu	Leu	Phe	Leu	Gly	Thr	Leu	Gly	Leu	
1				5					10					15		
Cys	Leu	Ala	Ala	Pro	Lys	Lys	Gly	Val	Arg	Trp	Cys	Val	Ile	Ser	Thr	
			20					25					30			
Ala	Glu	Tyr	Ser	Lys	Cys	Arg	Gln	Trp	Gln	Ser	Lys	Ile	Arg	Arg	Thr	
		35					40					45				
Asn	Pro	Met	Phe	Cys	Ile	Arg	Arg	Ala	Ser	Pro	Thr	Asp	Cys	Ile	Arg	
	50					55					60					
Ala	Ile	Ala	Ala	Lys	Arg	Ala	Asp	Ala	Val	Thr	Leu	Asp	Gly	Gly	Leu	
65					70					75					80	
Val	Phe	Glu	Ala	Asp	Gln	Tyr	Lys	Leu	Arg	Pro	Val	Ala	Ala	Glu	Ile	
				85					90					95		
Tyr	Gly	Thr	Glu	Glu	Asn	Pro	Gln	Thr	Tyr	Tyr	Tyr	Ala	Val	Ala	Val	
			100					105					110			
Val	Lys	Lys	Gly	Phe	Asn	Phe	Gln	Asn	Gln	Leu	Gln	Gly	Arg	Lys	Ser	
		115					120					125				
Cys	His	Thr	Gly	Leu	Gly	Arg	Ser	Ala	Gly	Trp	Asn	Ile	Pro	Ile	Gly	
	130					135					140					
Leu	Leu	Arg	Arg	Phe	Leu	Asp	Trp	Ala	Gly	Pro	Pro	Glu	Pro	Leu	Gln	
145					150					155					160	
Lys	Ala	Val	Ala	Lys	Phe	Phe	Ser	Gln	Ser	Cys	Val	Pro	Cys	Ala	Asp	
				165					170					175		
Gly	Asn	Ala	Tyr	Pro	Asn	Leu	Cys	Gln	Leu	Cys	Ile	Gly	Lys	Gly	Lys	
		180						185					190			
Asp	Lys	Cys	Ala	Cys	Ser	Ser	Gln	Glu	Pro	Tyr	Phe	Gly	Tyr	Ser	Gly	
		195					200					205				
Ala	Phe	Asn	Cys	Leu	His	Lys	Gly	Ile	Gly	Asp	Val	Ala	Phe	Val	Lys	
	210					215					220					
Glu	Ser	Thr	Val	Phe	Glu	Asn	Leu	Pro	Gln	Lys	Ala	Asp	Arg	Asp	Lys	
225					230					235					240	

Tyr Glu Leu Leu Cys Pro Asp Asn Thr Arg Lys Pro Val Glu Ala Phe  
 245 250 255  
 Arg Glu Cys His Leu Ala Arg Val Pro Ser His Ala Val Val Ala Arg  
 260 265 270  
 Ser Val Asn Gly Lys Glu Asn Ser Ile Trp Glu Leu Leu Tyr Gln Ser  
 275 280 285  
 Gln Lys Lys Phe Gly Lys Ser Asn Pro Gln Glu Phe Gln Leu Phe Gly  
 290 295 300  
 Ser Pro Gly Gln Gln Lys Asp Leu Leu Phe Arg Asp Ala Thr Ile Gly  
 305 310 315 320  
 Phe Leu Lys Ile Pro Ser Lys Ile Asp Ser Lys Leu Tyr Leu Gly Leu  
 325 330 335  
 Pro Tyr Leu Thr Ala Ile Gln Gly Leu Arg Glu Thr Ala Ala Glu Val  
 340 345 350  
 Glu Ala Arg Gln Ala Lys Val Val Trp Cys Ala Val Gly Pro Glu Glu  
 355 360 365  
 Leu Arg Lys Cys Arg Gln Trp Ser Ser Gln Ser Ser Gln Asn Leu Asn  
 370 375 380  
 Cys Ser Leu Ala Ser Thr Thr Glu Asp Cys Ile Val Gln Val Leu Lys  
 385 390 395 400  
 Gly Glu Ala Asp Ala Met Ser Leu Asp Gly Gly Phe Ile Tyr Thr Ala  
 405 410 415  
 Gly Lys Cys Gly Leu Val Pro Val Leu Ala Glu Asn Gln Lys Ser Arg  
 420 425 430  
 Gln Ser Ser Ser Ser Asp Cys Val His Arg Pro Thr Gln Gly Tyr Phe  
 435 440 445  
 Ala Val Ala Val Val Arg Lys Ala Asn Gly Gly Ile Thr Trp Asn Ser  
 450 455 460  
 Val Arg Gly Thr Lys Ser Cys His Thr Ala Val Asp Arg Thr Ala Gly  
 465 470 475 480  
 Trp Asn Ile Pro Met Gly Leu Leu Val Asn Gln Thr Gly Ser Cys Lys  
 485 490 495  
 Phe Asp Glu Phe Phe Ser Gln Ser Cys Ala Pro Gly Ser Gln Pro Gly  
 500 505 510  
 Ser Asn Leu Cys Ala Leu Cys Val Gly Asn Asp Gln Gly Val Asp Lys  
 515 520 525  
 Cys Val Pro Asn Ser Asn Glu Arg Tyr Tyr Gly Tyr Thr Gly Ala Phe  
 530 535 540  
 Arg Cys Leu Ala Glu Asn Ala Gly Asp Val Ala Phe Val Lys Asp Val  
 545 550 555 560  
 Thr Val Leu Asp Asn Thr Asn Gly Gln Asn Thr Glu Glu Trp Ala Arg  
 565 570 575